

**O'Bryen, Barbara**

---

**From:** Goldberg, Jeanine  
**Sent:** Tuesday, April 02, 2002 7:34 AM  
**To:** O'Bryen, Barbara  
**Subject:** 09/762,724- MSG Pneumocystis carinii

Please place results on DISK.

1. Please search SEQ ID NO: 17-20, 23, 24
2. Please search SEQ ID NO: 13, positions 2821-3072
3. Please search SEQ ID NO: 3, positions 2758-3006
4. Please search SEQ ID NO: 1, positions 2894-3042
5. Please search SEQ ID NO: 5, positions 2845-3090

THANK YOU  
Jeanine

Jeanine Enewold Goldberg  
1634  
CM1--12D11  
Mailbox-- 12E12  
306-5817

Point of Contact:  
Barb O'Bryen  
Technical Information Specialist  
STIC CM1 6A05 308-4291

*BoB*  
*4-2-02*

SEQ ID NO: 13, nucleotides 2821-3072

RESULT 5

PMCMMSGI

LOCUS PMCMMSGI 3363 bp DNA PLN 26-SEP-1994

DEFINITION Pneumocystis carinii B-cell receptor (msgI) gene, 3' end.

ACCESSION L27092

VERSION L27092.1 GI:535706

KEYWORDS B-cell receptor.

SOURCE Pneumocystis carinii DNA.

ORGANISM Pneumocystis carinii

Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;  
Pneumocystidaceae; Pneumocystis.

REFERENCE 1 (bases 1 to 3363)

AUTHORS Garbe, T.R. and Stringer, J.R.

TITLE Molecular characterization of clustered variants of genes encoding  
major surface antigens of human Pneumocystis carinii

JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)

MEDLINE 94314421

FEATURES Location/Qualifiers

source

1. .3363

/organism="Pneumocystis carinii"

/db\_xref="taxon:4754"

mat\_peptide

152. .3241

/gene="msgI"

/standard\_name="major surface antigen"

/evidence=experimental

/product="B-cell receptor"

gene

152. .3244

/gene="msgI"

CDS

<152. .3244

/gene="msgI"

/codon\_start=1

/product="B-cell receptor"

/protein\_id="AAA21645.1"

/db\_xref="GI:535707"

/translation="VARAVKRQVAGVKNNEAEERLFALIMRADYKDESKCKNKIKEYC  
DGLKNASLTSEEVHKELKDFCKDGSQGKKCEELKKNVEAKCENNFKTKLEGLVKKDASG  
LTNDDCKENERQCLFLEGACPDVEDCSKLRNLCYQKKREGVAEEVLLRALRGDLGNK  
TECEKKIKDVCPKIGQESDELTLLCLDQKKCTNLMTARDKKCNTLEEDVKKALENKN  
NLLGKCLPLLEHATFTEGTAKKASQCTPNKDCEDYLPKCDELAEECGKKGIYIHPGP  
DFDPTKPEPTVAEDIGLEELYKKAEDGVHIGKPPVRDATAALLIIONPDPKIQANE  
KEKCKKVLENKCKELKKHEVLGDLNQNAAASQSGTKKCEELEKELANSTKILSEKIKN  
KHLSGSGETIPWYKLSTFLSDSCARLESDFYFAQDKDPLKKECKNVKAAACYKRGDL  
ARANKVLQENMRGLLRGSNQS WLKKFQQELVKVCEKLKEENKGSFSDNDELFLVCVQPA  
KAARLLTHDLRMTTIFLRQQLDQKRDFPTVKTAGIREKQDLGKGFQKEITWPCHTL  
EQQCNRLGTTIILKQVLLNEHKDTLKT HENCVTYLKECKNKWSRRGDDRFSEFVCFQN  
ATCKLMVKDVQDRCKIFKENIKVSEIVDFLKNNTNNITTLERNCPSWHTYCNRFSSNC  
PDFSKKNPCTKIKNNCKPFYERKALEDALKVELRGKLSDENKCTAALKGYCTLAGNVN  
NASVRSCLKDNTQGSNKKTDEKVVEELCKKLMEEVKEQCETLPAELKQPADDLEKDVK  
TYEELKEEAKKAMNKSSLVLSFVKKDGNNTPKNNSKSEDKNVVSNEKDTIKHVKILRR  
GVKDVLVTELEAKAFDLAAEVFGRYVDLKERCEKLTLD CGIKDDCDGLKGVCGKIKKK  
CRDLKPLEVKSHEIVTESTTTTTTTTTTTTVDPKATECKSLQTTDTWVTQTSTHTSTST  
ITSTITSKITLTSTRCKPTKCTTGDDAEDVKPSEGLRVSGWNVMRGAI VAMVISFMI  
"

BASE COUNT 1312 a 505 c 753 g 793 t

ORIGIN

Query Match 80.5%; Score 202.8; DB 8; Length 3363;  
Best Local Similarity 90.9%; Pred. No. 2.8e-47;  
Matches 229; Conservative 0; Mismatches 17; Indels 6; Gaps 1;

```
Qy      1 gaatgcaaatccttacagacaacagacacatgggttacacagacatcgacacacacaagc 60
      |||
Db 2999 GAATGCAAATCCTTACAGACAACAGACACATGGGTTACACAGACATCGACACACACAAGC 3058

Qy     61 acgtctactatcacatctaccatcacatcaaaaataacattgacatcaacgaggcgatgc 120
      || ||||| |||
Db 3059 ACATCTACCATCACATCTACGATTACATCAAAAATAACATTGACATCAACAAGGCGGTGC 3118

Qy    121 aaaccaaccaagtgtacgacaggagaggaagatgatgcaggagacgtgaaaccgagtgc 180
      ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 3119 AAACCAACCAAGTGTACGACAGG-----GGATGATGCAGAAGACGTGAAGCCAAGTGAA 3172

Qy    181 gggctgaggatgagtggggtggaatgtgatgaggggggtgatagtagcaatgggttatttcg 240
      || ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
Db 3173 GGCTTGAGGGTGAGCGGGTGGAATGTGATGAGGGGGGCAATAGTAGCAATGGTTATTTTCG 3232

Qy    241 ttcatgatttag 252
      |||
Db 3233 TTCATGATTTAG 3244
```

SEQ ID NO: 3, nucleotides 2758-3006

RESULT 7

PMCMSGI

LOCUS PMCMSGI 3363 bp DNA PLN 26-SEP-1994

DEFINITION Pneumocystis carinii B-cell receptor (msgI) gene, 3' end.

ACCESSION L27092

VERSION L27092.1 GI:535706

KEYWORDS B-cell receptor.

SOURCE Pneumocystis carinii DNA.

ORGANISM Pneumocystis carinii

Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;  
Pneumocystidaceae; Pneumocystis.

REFERENCE 1 (bases 1 to 3363)

AUTHORS Garbe,T.R. and Stringer,J.R.

TITLE Molecular characterization of clustered variants of genes encoding  
major surface antigens of human Pneumocystis carinii

JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)

MEDLINE 94314421

FEATURES Location/Qualifiers

source

1. .3363

/organism="Pneumocystis carinii"

/db\_xref="taxon:4754"

mat\_peptide

152. .3241

/gene="msgI"

/standard\_name="major surface antigen"

/evidence=experimental

/product="B-cell receptor"

gene

152. .3244

/gene="msgI"

CDS

<152. .3244

/gene="msgI"

/codon\_start=1

/product="B-cell receptor"

/protein\_id="AAA21645.1"

/db\_xref="GI:535707"

/translation="VARAVKRQVAGVKNNEAEERLFALIMRADYKDESKCKNKIKEYC  
DGLKNASLTSEEVHKELKDFCKDGSQGKKCEELKKNVEAKCNNFKTKLEGLVKKDASG  
LTNDDCKENERQCLFLEGACPDLVEDCSKLRNLCYQKKREGVAEEVLLRALRGDLGNK  
TECEKKIKDVCPKIGQESDELTLLCLDQKKTCTNLMTARDKKCNTLEEDVKKALENKN  
NLLGKCLPLLEHATFTEGTAKKASQCTPNKDCEDYLPKCELAEECGKKGIIYIHGPGP  
DFDPTKPEPTVAEDIGLEELYKKAEDGVHIGKPPVRDATAALLALLIQNPDPKIQANE  
KEKCKKVLENKCKELKKHEVLGDLNQNAAASQSGTKKCEELEKELANSTKILSEKIKN  
KHLSGSGETIPWYKLSTFLSDSDCARLESDFYFAQDKDPLKKECKNVKAACYKRGDL  
ARANKVLQENMRGLLRGSNSWLKKFQQELVKVCEKLKEENKGSFSNDELFLVCVQPA  
KAARLLTHDLRMTTIFLRQQLDQKRDFPTVKTAGIREKCQDLGKGFQKEITWPCHTL  
EQQCNRGLGTTEILKQVLLNEHKDTLKTENCVTYLKECKNKWSRRGDDRFSEVCFVQN  
ATCKLMVKDVQDRCKIFKENIKVSEIVDFLKNNTNNITTLERNCPSWHTYCNRFSSNC  
PDFSKKNPCTKIKNNCKPFYERKALEDALKVELRGKLSDENKCTAALKGYCTLAGNVN  
NASVRS�CKDNTQGSNKKTDKVVVEELCKKLMEEVKEQCETLPAELKQPADDLEKDVK  
TYEELKEEAKKAMNKSSLVLSFVKKDGNNTPKNNSKSEDKNVVSNEKDTIKHVKILRR  
GVKDVLVTELEAKAFDLAAEVFGRYVDLKERCEKLTLDGCIKDDCDGLKGVCCKIKKK  
CRDLKPLEVKSHEIVTESTTTTTTTTTTTTVDPKATECKSLQTTDTWVTQTSTHTSTST  
ITSTITSKITLTSTRCKPTKCTTGDDAEDVKPSEGLRVSGWNVMRGAIVAMVISFMI  
"

BASE COUNT 1312 a 505 c 753 g 793 t

ORIGIN

Query Match 61.6%; Score 153.4; DB 8; Length 3363;  
Best Local Similarity 79.1%; Pred. No. 1.2e-33;  
Matches 197; Conservative 0; Mismatches 46; Indels 6; Gaps 1;

```
Qy      1 gactgccactctttacagacaacagatacgtgggtcacaagacgtcgacccatactagc 60
      || ||| | || ||||| ||||| || ||||| ||| ||| ||| ||| |||
Db    2999 GAATGCAAATCCTTACAGACAACAGACACATGGGTTACACAGACATCGACACACACAAGC 3058

Qy      61 acatccacaaccacatctacagtcacgtcaagaataacgttgacctcgacaagacggtgt 120
      ||||| || | ||||| || | || ||| ||||| ||| ||||| |||||
Db    3059 ACATCTACCATCACATCTACGATTACATCAAAAATAACATTGACATCAACAAGGCGGTGC 3118

Qy     121 aagcctacgaagtgtacgacaggagaggaagatgaagcaggagacgtgaaaccgagtgaa 180
      || || || ||||| ||||| ||||| ||||| ||||| ||| |||||
Db    3119 AAACCAACCAAGTGTACGACAGG-----GGATGATGCAGAAGACGTGAAGCCAAGTGAA 3172

Qy     181 gggttgaggatgagtggtgagtggtgatgaggggggtgttattagcaatgacgatttca 240
      || ||||| |||| || |||| ||||| ||||| || ||||| |||||
Db    3173 GGCTTGAGGGTGAGCGGGTGAATGTGATGAGGGGGGCAATAGTAGCAATGGTTATTTTCG 3232

Qy     241 ttcattgatt 249
      ||||| |||
Db    3233 TTCATGATT 3241
```

SEQ ID NO: 1, nucleotides 2894-3042

RESULT 7

PMCMGSI

LOCUS PMCMGSI 3363 bp DNA PLN 26-SEP-1994

DEFINITION Pneumocystis carinii B-cell receptor (msgI) gene, 3' end.

ACCESSION L27092

VERSION L27092.1 GI:535706

KEYWORDS B-cell receptor.

SOURCE Pneumocystis carinii DNA.

ORGANISM Pneumocystis carinii

Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;

Pneumocystidaceae; Pneumocystis.

REFERENCE 1 (bases 1 to 3363)

AUTHORS Garbe,T.R. and Stringer,J.R.

TITLE Molecular characterization of clustered variants of genes encoding major surface antigens of human Pneumocystis carinii

JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)

MEDLINE 94314421

FEATURES Location/Qualifiers

source

1. .3363

/organism="Pneumocystis carinii"

/db\_xref="taxon:4754"

mat\_peptide

152. .3241

/gene="msgI"

/standard\_name="major surface antigen"

/evidence=experimental

/product="B-cell receptor"

gene

152. .3244

/gene="msgI"

CDS

<152. .3244

/gene="msgI"

/codon\_start=1

/product="B-cell receptor"

/protein\_id="AAA21645.1"

/db\_xref="GI:535707"

/translation="VARAVKRQVAGVKNNEAEERLFALIMRADYKDESKCKNKIKEYC  
DGLKNASLTSEEVHKELKDFCKDGSQGKKCEELKKNVEAKCENNFKTKLEGLVKKDASG  
LTNDDCKENERQCLFLEGACPDLEDSCSKLRNLCYQKKREGVAEEVLLRALRGDLGNK  
TECEKKIKDVCPKIGQESDELTLLCLDQKKTCTNLMTARDKKCNTLEEDVKKALENKN  
NLLGKCLPLLEHATFTEGTAKKASQCTPNKDCEDYLPKCELAEECGKKGIIYIHPGP  
DFDPTKPEPTVAEDIGLEELYKKAEDGVHIGKPPVRDATAALLIIONPDPKIQANE  
KEKCKKVLENKCKELKKHEVLGDLNQNAAASQSGTKKCEELEKELANSTKILSEKIKN  
KHLSGSGETIPWYKLSTFLSDSCARLESDFYFAQDKDPLKKECKNVKAAACYKRGDL  
ARANKVLQENMRGLLRGSNSWLKKFQQELVKVCEKLKEENKGSFSDNDELFLVCVQPA  
KAARLLTHDLRMTTIFLRQQLDQKRDFPTVKTAGIREKCQDLGKGFQKEITWPCHTL  
EQQCNRLLGTTTEILKQVLLNEHKDTLKTHENCVTYLKECKNKWSRRGDDRFSEVCFQON  
ATCKLMVKDVQDRCKIFKENIKVSEIVDFLKNNTNNITTLERNCPSWHTYCNRFSSNC  
PDFSKKNPCTKIKNNCKPFYERKALEDALKVELRGKLSDENKCTAALKGYCTLAGNVN  
NASVRSICKDNTQGSNKKTDKVVVEELCKKLMEEVKEQCETLPAELKQPADDLEKDVK  
TYEELKEEAKKAMNKSSLVLSFVKKDGNNTPKNNSKSEDKNVVSNEKDTIKHVKILRR  
GVKDVLVTELEAKAFDLAAEVFGRYVDLKERCEKLTLDGCIKDDCDGLKGVCGKIKKK  
CRDLKPLEVKSHEIVTESTTTTTTTTTTTTVDPKATECKSLQTTDTWVTQTSTHTSTST  
ITSTITSKITLTSTRCKPTKCTTGDDAEDVKPSEGLRVSGWNVMRGAIVAMVISFMI  
"

BASE COUNT 1312 a 505 c 753 g 793 t

ORIGIN



SEQ ID NO: 5, 2845-3090 nucleotides

RESULT 1

AF033208

LOCUS AF033208 3089 bp DNA PLN 10-SEP-1998

DEFINITION Pneumocystis carinii f. sp. hominis clone HUMSG11 major surface glycoprotein (MSG) gene, partial cds.

ACCESSION AF033208

VERSION AF033208.1 GI:3560512

KEYWORDS .

SOURCE Pneumocystis carinii f. sp. hominis.

ORGANISM Pneumocystis carinii f. sp. hominis  
Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;  
Pneumocystidaceae; Pneumocystis.

REFERENCE 1 (bases 1 to 3089)

AUTHORS Garbe,T.R. and Stringer,J.R.

TITLE Molecular characterization of clustered variants of genes encoding major surface antigens of human Pneumocystis carinii

JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)

MEDLINE 94314421

REFERENCE 2 (bases 1 to 3089)

AUTHORS Mei,Q., Turner,R.E., Sorial,V., Klivington,D., Angus,C.W. and Kovacs,J.A.

TITLE Characterization of major surface glycoprotein genes of human Pneumocystis carinii and high-level expression of a conserved region

JOURNAL Infect. Immun. 66 (9), 4268-4273 (1998)

MEDLINE 98380374

REFERENCE 3 (bases 1 to 3089)

AUTHORS Mei,Q., Turner,R., Sorial,V., Klivington,D., Angus,C.W. and Kovacs,J.A.

TITLE Direct Submission

JOURNAL Submitted (07-NOV-1997) Critical Care Medicine Dept., National Institutes of Health, 10 Center Drive, MSC 1662, Bethesda, MD 20892-1662, USA

FEATURES Location/Qualifiers

source

1. .3089  
/organism="Pneumocystis carinii f. sp. hominis"  
/sub\_species="hominis"  
/specific\_host="Homo sapiens"  
/db\_xref="taxon:42068"  
/clone="HUMSG11"  
/note="derived from HIV-infected human with P. carinii pneumonia"

gene

<1. .3089

/gene="MSG"

CDS

<1. .3089

/gene="MSG"

/note="surface antigen"

/codon\_start=3

/product="major surface glycoprotein"

/protein\_id="AAC34971.1"

/db\_xref="GI:3560513"

/translation="ARAVKRRAKGAQNSIDEEHVLALILKKNGLDTCKTKLEEYCK  
TLTNAGLNPEKVHEKLDKDFCDNGKRNEKCQDLKNVQKCIKFQGLQTAAGKKISEL  
TDEDCKKNEQQCLFLEGACPTTELKDDCNKLRNNCYQKERNNVAAEEVLLRALRGDLNET  
KTCEKKLKEVCPLERESDELTELCLYQKTTVCVSLVTGKSKCDTLEKEVEEALKKNE  
LREKCLLLLEQCYFHRGNCEGDKSKCNKPNKDKKEYVPECDELAEKCGKENIVMHP  
GSDFDPTKPEPTLAEDIGLEELYKRAEEDGIFVGRQHVRRDATAALLALLKKTLLKKEEC  
IKALKKNCENPHEHEALENLCKENKPSSDGTKKCDELEKDVNKTCTSLTSTILKNRLY  
ISPDGIAEWGKLPTFLSDEDCAKLESYCFYYKETCPDVKEACMNVRAACYKRGLDARA  
NSVLQKNMRGLLHGSNKDWLKKFQQELAKVCEKLGKNGSFSNDEL FVLCIQPAKAAR  
LLTHHHQMRVI FLRQQLDQKRDFPTDKCKELGRKCQDLGKDSKEITWPCHTLEQQCN  
RLGITEILKQILLDEHKDTLKSHENCAKYLKRKCHKWSRRGDDRFVFCVFNATCEL





SEQ ID NO: 5, nucleotides 2845-3090

RESULT 6

PMCMGSI

LOCUS PMCMGSI 3363 bp DNA PLN 26-SEP-1994

DEFINITION Pneumocystis carinii B-cell receptor (msgI) gene, 3' end.

ACCESSION L27092

VERSION L27092.1 GI:535706

KEYWORDS B-cell receptor.

SOURCE Pneumocystis carinii DNA.

ORGANISM Pneumocystis carinii

Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;  
Pneumocystidaceae; Pneumocystis.

REFERENCE 1 (bases 1 to 3363)

AUTHORS Garbe,T.R. and Stringer,J.R.

TITLE Molecular characterization of clustered variants of genes encoding  
major surface antigens of human Pneumocystis carinii

JOURNAL Infect. Immun. 62 (8), 3092-3101 (1994)

MEDLINE 94314421

FEATURES Location/Qualifiers

source

1. .3363

/organism="Pneumocystis carinii"

/db\_xref="taxon:4754"

mat\_peptide

152. .3241

/gene="msgI"

/standard\_name="major surface antigen"

/evidence=experimental

/product="B-cell receptor"

gene

152. .3244

/gene="msgI"

CDS

<152. .3244

/gene="msgI"

/codon\_start=1

/product="B-cell receptor"

/protein\_id="AAA21645.1"

/db\_xref="GI:535707"

/translation="VARAVKRQVAGVKNNEAEERLFALIMRADYKDESKCKNKIKEYC  
DGLKNASLTSEEVHKELKDFCKDGSQGKKCEELKKNVEAKCNNFKTKLEGLVKKDASG  
LTNDDCKENERQCLFLEGACPDLEDCKSLRNLCYQKKREGVAEEVLLRALRGDLGNK  
TECEKKIKDVCPKIGQESDELTLLCLDQKKCTNLMTARDKKCNTLEEDVKKALENKN  
NLLGKCLP LLEHATFTEGTAKKASQCTPNKDCEDYLPKCELAEECGKKGIYIHPGP  
DFDPTKPEPTVAEDIGLEELYKKAEDGVHIGKPPVRDATAALLALIQNPDPKIQANE  
KEKCKKVLENCKELKKHEVLGDLNQNAAASQSGTKKCEELEKELANSTKILSEKIKN  
KHLSGSGETIPWYKLSTFLSDSDCARLESDFYFAQDKDPLKKECKNVKAAACYKRGD  
ARANKVLQENMRGLLRGSNQSWLKKFQQELVKVCEKLKEENKGSFSDNDELFLVLCVQPA  
KAARLLTHDLRMTTIFLRQQLDQKRDFPTVKTAKGIREKQDLGKGFKQKEITWPCHTL  
EQQCNRGLGTTEILKQVLLNEHKDTLKTENCVTYLKEKCNKWSRRGDDRFSEVCFVQN  
ATCKLMVKDVQDRCKIFKENIKVSEIVDFLKNNTNITTLEARNCPSWHTYCNRFSSNC  
PDFSKKNPCTKIKNNCKPFYERKALEDALKVELRGKLSDENKCTAALKGYCTLAGNVN  
NASVRS LCKDNTQGSNKKTDKVVVEELCKKLMEEVKEQCETLPAELKQPADDLEKDVK  
TYEELKEEAKKAMNKSSLVLSFVKKDGNNTPKNNSKSEDKNVVSNKEDTIKHVKILRR  
GVKDVLVTELEAKAFDLAAEVFGRYVDLKERCEKLTLDGCIKDDCDGLKGVCGKIKKK  
CRDLKPLEVKSHEIVTESTTTTTTTTTTTTTVTDPKATECKSLQTTDTWVTQTSTHTSTST  
ITSTITSKITLTSTRCKPTKCTTGDDAEDVKPSEGLRVSGWNVMRGAIVAMVISFMI  
"

BASE COUNT 1312 a 505 c 753 g 793 t

ORIGIN

Query Match 87.6%; Score 215.6; DB 8; Length 3363;  
Best Local Similarity 92.3%; Pred. No. 3.4e-49;  
Matches 227; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy	1	gaatgcaaatccttacagacaacagatacatgggttacacagacatcgacacacacaagc	60
Db	2999	GAATGCAAATCCTTACAGACAACAGACACATGGGTTACACAGACATCGACACACACAAGC	3058
Qy	61	acgtctaccatcacatctaccatcacatcaaaaataacattgacatcaacgaggcgatgc	120
Db	3059	ACATCTACCATCACATCTACGATTACATCAAAAATAACATTGACATCAACAAGGCGGTGC	3118
Qy	121	aaaccaaccaagtgtacgacaggggatgaagcaggagacgtgaaaccgagtgagggattg	180
Db	3119	AAACCAACCAAGTGTACGACAGGGGATGATGCAGAAGACGTGAAGCCAAGTGAAGGCTTG	3178
Qy	181	aagatgagtggggtggagcgtgatgaggggggtgatagtagcaatggttatttcgttcatg	240
Db	3179	AGGGTGAGCGGGTGGGAATGTGATGAGGGGGGCAATAGTAGCAATGGTTATTTCGTTCATG	3238
Qy	241	atttag	246
Db	3239	ATTTAG	3244

LOCUS PCU39660 629 bp mRNA P. 29-NOV-1995  
DEFINITION **Pneumocystis carinii** major surface glycoprotein (msg) mRNA, partial  
cds.  
ACCESSION U39660  
NID g1079706  
KEYWORDS .  
SOURCE **Pneumocystis carinii**.  
ORGANISM **Pneumocystis carinii**  
Eukaryotae; mitochondrial eukaryotes; Fungi; Ascomycota;  
Archaeascomycetes; Pneumocystidaceae; Pneumocystis.  
REFERENCE 1 (bases 1 to 629)  
AUTHORS Kovacs, J.A., Edman, J.C. and Angus, C.W.  
TITLE Direct Submission  
JOURNAL Submitted (30-OCT-1995) J.A. Kovacs, CCMD, NIH, Building 10, Rm  
7D43, MSC 1662, Bethesda, MD 20892-1662, USA  
FEATURES  
source Location/Qualifiers  
1..629  
/organism="Pneumocystis carinii"  
/db\_xref="taxon:4754"  
gene 55..629  
/gene="msg"  
CDS 55..>629  
/gene="msg"  
/codon\_start=1  
/product="major surface glycoprotein"  
/db\_xref="PID:g1079707"  
/translation="MRIAFFALFAQLSCILVYSIAERDFMSLDEIYGGDISFDHEKLE  
FNEYNQVLQMLEKAKKLGTGFVDRTKDFSNNRYEGRIELNHLGRRPGVDYFRKGGDVF  
TDGYPRGGHLIEDELSEEAAMARPVKRQAVQGAQDEIDEKHLALFIVKDKYKEEQKCK  
EELEKYCKELKEADKNLENVDDKVKGLCDDK"  
misc\_difference 228..234  
/gene="msg"  
/note="in some cDNA clones, there are 6 A's at this  
position, instead of 7. This appears to be due to an error  
in reverse transcription"  
/replace="aaaaaa"  
BASE COUNT 218 a 73 c 154 g 184 t  
ORIGIN  
1 ggaaaatata ttttttcttg atatccgtct cgtcttcagt ttgtttgtgc aataatgagg  
61 attgcatttt ttgcgctttt tgcgcaactt agttgtattt tagtttattc aatagcagaa  
121 agggatttca tgtcattaga tgaaatatat ggaggcgata taagttttga tcatgaaaaa  
181 ctcgaattta acgaatataa tcaagtttta caaatgcttg aaaaggcaaa aaaattggga  
241 accggctttg ttgatagaac caaagatttt tctaatagac gatatgaagg gagaattgag  
301 ttaaatacatt tggggagacg cccaggagtc gactatttta ggaaaggtgg ggatgttttt  
361 actgatggtt atcctcgtgg aggtcatttg atcgaggatg agttgtccga agaggcggca  
421 atggcacggc cggttaagag gcaagcagta caaggagcac aagatgagat tgatgagaaa  
481 caccttttgg ctttcattgt gaaggacaaa tataaagaag aacaaaaatg caaagaagaa  
541 ctcgagaaat attgtaaaga gttgaaggaa gcagataaaa atctagagaa tgtggatgat  
601 aaagttaaag gactttgtga tgataaaaa  
//

LOCUS AF043102 4002 bp mRNA PLN 06-JUN-1998  
DEFINITION *Pneumocystis carinii* surface glycoprotein A (gpA) mRNA, complete cds.  
ACCESSION AF043102  
NID g3184385  
KEYWORDS .  
SOURCE *Pneumocystis carinii*.  
ORGANISM *Pneumocystis carinii*  
Eukaryota; Fungi; Ascomycota; Archiascomycetes; Pneumocystidaceae; *Pneumocystis*.  
REFERENCE 1 (bases 1 to 4002)  
AUTHORS Haidaris, C.G., Medzihradsky, O.F., Gigliotti, F. and Simpson-Haidaris, P.J.  
TITLE Molecular characterization of mouse *Pneumocystis carinii* surface glycoprotein A  
JOURNAL DNA Res. 5, 77-85 (1998)  
REFERENCE 2 (bases 1 to 4002)  
AUTHORS Haidaris, C.G., Medzihradsky, O.F., Gigliotti, F. and Simpson-Haidaris, P.J.  
TITLE Direct Submission  
JOURNAL Submitted (14-JAN-1998) Microbiol/Immunol, U. Rochester Med. Ctr., Box 672, 601 Elmwood Ave., Rochester, NY 14642, USA  
FEATURES  
source Location/Qualifiers  
1..4002  
/organism="Pneumocystis carinii"  
/db\_xref="taxon:4754"  
/lab\_host="C.B-17 SCID mouse"  
gene 1..4002  
/gene="gpA"  
CDS 52..3849  
/gene="gpA"  
/codon\_start=1  
/product="surface glycoprotein A"  
/db\_xref="PID:g3184386"  
/translation="MRIAFFALFAQLSYVLGSFLRDRDLPNEEDVYGYENFGLDPNSP  
EPSEFLNIIEMLRNAKRLGSGKINQQGFLSNRKARRDFDLRSCNRPGVDYFRKSDYD  
GFFSEDFSSSEDYSQDKRWVEEVAQKEAAMAQPVKRQAAGQAAGNDEIKEEQVLGLIVK  
SGYNNDNKCKANLKHCEELKKIDGKLESVDVKVKGKLCENGKEGKCKELKKKLETEL  
GAFKTEVENALNNLTDEKCRKYEEKCLLLEEGDPNNLEEKCVKLDRDCYRQRQGVAK  
EILLRALEGKVNNDCKKRMKEICQGLSEYSDELVFSCFNSDKTCEYLQKNHGDSCK  
PLEKELEDKELVEKCEYLEKCYFYGSSCKDTKCDKVNCKKGKIEYEGPKLDFSPV  
REKPRFPEKIEVENLYKKEEAKGIIVGKPKYKTLRDLALLLIKERNKGDEGEKCKKAL  
EDCESFKHLDYGLEELCGDKDKEDRCKELVEVEDRCTNFKLEYLKGKSTEFKDKES  
DYFSWGQVSKLVSMEDCIKFESECFHLERVCTNKIGKACENVRVACYKKGQDRVLNRY  
FQEGKGLIGDLELVTENLEKQKQSVVGNVYTKLKEDRRYFTKCHLPTKLCYELDDVI  
LQSEELVVLNLRDFFRKEDCVLKKKCDLESYSYLNHEKCDTLNRRCEYLKVTET  
LRKRLKRGDDALRTQGNCTAVLKKECEELSKRGKDEFVSVCALREETCSFMVEQTEN  
ECLFLKNNMENGKIINKIEKGNETLVEELCTLFDPYCHQYIENCPDRLKKASNSNKG  
VCLELEEKCKPFFEKLLENELTHELKGSLDKDDECKKALRKHCSEQKNSANQKFNSL  
CNTDKDKDVEEKVCKKLVEKVKRKCPTLENKLNEEKNELKKKKDEYEKAKQSESEKFTK  
EAKLVLSRPEQDGGGGSKAQDGSVPKPVGPPVQPPAPAQPTPGGVPAPTLPAPQPT  
SGGAPLPVPPAAPAPGAPSTPGTAAAPAGPAAGTPTSTPTPPAGPAGSGGTTPGAP  
AGPPAPGGSTPSGTTNTSNVILVRRFTVSGEVSEPEKKAFVATARALELYLELKEKCK  
GLKGDCEFRKDCPKCETVCKEIDELCEGIEGLKVTPHHTVTSTATQTTTTTATTTTTT  
TTTTTTTTTTTTTTTATTTESVDGKVTETCTLVOTTDVTSTSLHTSTLTSTSTVTST  
VTLTSMRKCKPTRCTSDSSKETETQKEEKEEEVKPNEGKIRVPEMIKIMLLGVIVM  
GML"  
BASE COUNT 1616 a 487 c 940 g 959 t  
ORIGIN  
1 caaaaatata tttttcttga tatccgtctc ttttctcaaa tttgtgcaat tatgaggatt  
61 gcattttttg cgctttttgc gcaacttagt tatgtcttgg ggagtttttt aagagataga  
121 gattttgcaa atgaggaaga tgtttatggt tatgaaaatt ttgggttgga tcccaattca  
181 ccggaaccga gtgaattttt aaatattata gaaatgcttc gaaatgcaa gcgatttagga  
241 agtggaaga ttaaccaggg tcaatttctc tcgaatcgta aggctaggag agattttgat

LOCUS D31920 81 bp mRNA PL 22-SEP-1997  
DEFINITION **Pneumocystis carinii** MSG mRNA for major surface glycoprotein,  
5'UTR.  
ACCESSION D31920  
NID g2443384  
KEYWORDS MSG.  
SOURCE **Pneumocystis carinii** cDNA to mRNA.  
ORGANISM **Pneumocystis carinii**  
Eukaryotae; Fungi; Ascomycota; Archaeascomycetes;  
Pneumocystidaceae; **Pneumocystis**.  
REFERENCE 1 (sites)  
AUTHORS Wada,M., Sunkin,S.M., Stringer,J.R. and Nakamura,Y.  
TITLE Antigenic variation by positional control of major surface  
glycoprotein gene expression in **Pneumocystis carinii**  
J. Infect. Dis. 171 (6), 1563-1568 (1995)  
MEDLINE 95287050  
REFERENCE 2 (bases 1 to 81)  
AUTHORS Nakamura,Y.  
TITLE Direct Submission  
JOURNAL Submitted (24-JUN-1994) to the DDBJ/EMBL/GenBank databases.  
Yoshikazu Nakamura, Institute of Medical Science, University of  
Tokyo, Tumor Biology; Shirokanedai 4-6-1, Minato-ku, Tokyo 108,  
Japan (E-mail:nak@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5307,  
Fax:03-5449-5415)  
FEATURES Location/Qualifiers  
source 1..81  
/organism="Pneumocystis carinii"  
/db\_xref="taxon:4754"  
gene 1..81  
/gene="MSG"  
5'UTR 1..81  
/gene="MSG"  
/note="major surface glycoprotein"  
BASE COUNT 30 a 10 c 13 g 28 t  
ORIGIN  
1 attcatgaag taactatcaa gcttaataaa ggaccaata taggagtata ttgttaatat  
61 ttatgctata tattgggata c  
//

LOCUS D31914 496 bp mRNA PL 22-SEP-1997  
 DEFINITION *Pneumocystis carinii* MSG mRNA for major surface glycoprotein,  
 5'UTR.  
 ACCESSION D31914  
 NID g2443378  
 KEYWORDS MSG.  
 SOURCE *Pneumocystis carinii* cDNA to mRNA.  
 ORGANISM *Pneumocystis carinii*  
 Eukaryotae; Fungi; Ascomycota; Archaeascomycetes;  
 Pneumocystidaceae; *Pneumocystis*.  
 REFERENCE 1 (sites)  
 AUTHORS Wada,M., Sunkin,S.M., Stringer,J.R. and Nakamura,Y.  
 TITLE Antigenic variation by positional control of major surface  
 glycoprotein gene expression in *Pneumocystis carinii*  
 J. Infect. Dis. 171 (6), 1563-1568 (1995)  
 JOURNAL  
 MEDLINE 95287050  
 REFERENCE 2 (bases 1 to 496)  
 AUTHORS Nakamura,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-JUN-1994) to the DDBJ/EMBL/GenBank databases.  
 Yoshikazu Nakamura, Institute of Medical Science, University of  
 Tokyo, Tumor Biology; Shirokanedai 4-6-1, Minato-ku, Tokyo 108,  
 Japan (E-mail:nak@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5307,  
 Fax:03-5449-5415)  
 FEATURES Location/Qualifiers  
 source 1..496  
 /organism="*Pneumocystis carinii*"  
 /db\_xref="taxon:4754"  
 gene 1..496  
 /gene="MSG"  
 5'UTR 1..496  
 /gene="MSG"  
 /note="major surface glycoprotein"  
 BASE COUNT 214 a 50 c 89 g 143 t  
 ORIGIN  
 1 caactatcaa gcttaataaa ggacccaata taggagtata ttgttaatat ttatgctata  
 61 tattgggatc cagaaataat atgaaatgct tttttccaaa attaataaaa ttatagttct  
 121 aagaatatta attttaattg ttcctactta tggcaatcaa agcacacatt tgaaagccag  
 181 agattctcaa gtaagcgag caaatactcc gttaaggat attgttgta aggatgaaga  
 241 aatccttgct tatattttga gagaagatta taaagataaa tgcgaactaa agcttgagga  
 301 atattgtaaa gaattgagg atatatagatca aggattaaat aaagttcata ctatagttaa  
 361 agaaatttgt gatgatgaaa aacgagacgg aaaatgcaaa gaactgaaag acaaagttaa  
 421 aaaggaattg gaaactttta aagaggaact tgaaaaagca ttgaaagaca taaaagatga  
 481 aaattgtgaa aaatat  
 //

LOCUS AF035226 1945 bp mRNA PL 02-JAN-1998  
 DEFINITION *Pneumocystis carinii* glycoprotein A (gpA) mRNA, partial cds.  
 ACCESSION AF035226  
 NID g2739185  
 KEYWORDS  
 SOURCE *Pneumocystis carinii*.  
 ORGANISM *Pneumocystis carinii*  
 Eukaryotae; Fungi; Ascomycota; Archiascomycetes; Pneumocystidaceae;  
*Pneumocystis*.  
 REFERENCE 1 (bases 1 to 1945)  
 AUTHORS Guadiz, G., Haidaris, C.G., Maine, G.N. and Simpson-Haidaris, P.J.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-NOV-1997) Medicine-Vascular Medicine, University of  
 Rochester, 601 Elmwood Avenue, Box 610, Rochester, NY 14642, USA

FEATURES  
 source Location/Qualifiers  
 1..1945  
 /organism="Pneumocystis carinii"  
 /db\_xref="taxon:4754"  
 /tissue\_type="infected lung"  
 /clone="2-1A"  
 gene <1..1945  
 /gene="gpA"  
 CDS <1..1914  
 /gene="gpA"  
 /codon\_start=1  
 /product="glycoprotein A"  
 /db\_xref="PID:g2739186"  
 /translation="KKELRKKNFNCDDSCSNEKDNWEKLVKLRDEDCAKLQSKCFYLE  
 KHCSANLKKPCNNVRVTCYKRGKAAAMTLLSVLKGKLPKPGPNNDYKDCQKALLEKC  
 KEVRNISAIEVFEMCLYPENTCQNLSDVQQKANYLSLFSLQNRDHPSQEDCVLQEK  
 EALEADADWLRPPCETLRTHCNFLYLSEKLRHLLSEGKGLSSNEICNKLNERCHS  
 WHKKKNETYAFPCALRNESCELMVVRVGKHCDEFKENLKNYNTTLIKNPSEETCLLWA  
 PHCENLTPNCEEKVAKDINETCRKLQEKCSVLKRNKLNKRLKRELKGLTDGKKTEC  
 IDYFKSLCTQADHPNKTALDILCKENGQNIIDDDKAKEKKCQPLIDEVQWECPLLKTKL  
 EKAKDEVAKKADEYKKLKEEAETETETKLNVTLPKEAGKATLIRDPKGSPPAKRS  
 VKVEVSVSPITKEHAEIDLDSQALDLYFELREEDDALILDCDFKEDCPQCKDSCNSI  
 LATCKGLQPLTIAEPRIIEDQKPEAGKEDTPEGGQAGGQAAGKECTSITTSDTWVTST  
 STHTTTTTSTKTTTTVKVTLTSTRCKPTKCTTGKGEPTSRGDEEEVVPSSGRKWLPS  
 LSIVIVGVVVALV"

BASE COUNT 744 a 279 c 451 g 471 t

## ORIGIN

```

1 aaaaaagaac ttcgtaaaaa gaatttctgc aatgacgata gctgcagtaa tgaaaaagat
61 aattgggaga aattagtaaa gttaagagat gaagattgtg caaaattaca atctaagtgt
121 ttttatttag aaaaacattg ttcagcta attaaaaaac catgtaacaa tgttagagta
181 acctgttaca agagaggatt aaaggcagcg gctatgactt tgttagagag tgtgtgaaa
241 gggaaactta aaccaggtcc aaataatgat tataaagatt gccaaaaagc attattagaa
301 aagtgtaaag aagtgtgaaa tatcagtgca gaagtgtttg agatgtgttt atatccagaa
361 aatacgtgtc aaaatctttc tattgatgtt caacaaaaag caaattattt aagccttttt
421 tcacttcaaa acagggacca tccaagtcaa gaagactgtg ttgaactaca agagaaatgt
481 gaagctttag aagcagatgc agactggcct cgacctcctt gtgaaacatt gagaacgcac
541 tgcaactttc ttacctttc agaaaagttg aaacatcatt tgttgagtga agggaaaggt
601 aaattaagta gcaatgagat ttgcaacaaa gagttgaatg agaggtgtca ctcatggcat
661 aaaaaaaaga acgaaacgta tgcttttccg tgtgctttgc gcaatgaatc ttgtgagttg
721 atgggttggc gtgtaggaaa acattgtgat gaattttaaag aaaacttgaa aaattataat
781 acaacattaa ttaaaaatcc atcagaagaa acatgtcttt tgtgggcacc acactgtgaa
841 aatctgacgc caaactgtga agaaaagggt gctaaagata tcaatgagac ttgccgtaaa
901 ctccaggaga aatgttcacc agttcttgag aagaggaatt tgaaaaataa attgaaacgt
961 gagttgaaaag ggaaattaac ggatggaaaag aagacagaat gtatcgatta ttttaaaagt
1021 ctttgtacac aggcagatca tccaaataaa acggcacttg acatattatg taaagaaaat
1081 ggacaaaata ttgatgatga taaagataaa gaaaagaaat gccaacctac cattgatgaa
1141 gtacaatggg agtgtccact attaaaaaca aagttagaaa aagcaaagga tgaagtggca
1201 aagaaggcag atgaatataa aaaactaaaa gaagaggcag aaaaggctat tacggaaaca
1261 aaactcaatg ttacagaact accaaaggaa gcaggaaaag caaccctgat tagaagagat
1321 ccaaaaggca gcagtccagc caagagatca gtaaaaggtag aggttgtgtc tgttccaata

```



```
1381 acgaaagaac at      aggc aattgattta ttatctcaag cat      ttct ttatttcgag
1441 ttgagagagg agtg      gatgc gttaatatta gattgtgact tta      agga ttgtcctcaa
1501 tgtaaagatt catgtaatag tatacttgct acctgtaaag ggttacaacc tcttacaatc
1561 gcggaaccac gtattattga agatcagaaa ccagaagctg gaaaggaaga tactcctgaa
1621 ggaggacaag ctggaggaca agctgcggga aaggagtgc cctctatcac aacaagtgac
1681 acatgggtca cgagtacgtc tacgcatacg acgacgacca cctcgacaaa gacgacgaca
1741 gttaaagtga cgttaacgtc aacgaggaag tgtaagccga ccaagtgtac cactggaaaa
1801 ggagagccga caagtagagg tgatgaagag gaggaggtgg tgccgagtgg aggaaggaag
1861 tggttgccaa gtctaagtat agttattgtt ggtgtggtgg tggcactcgt gtagatcatg
1921 atgaataaga cgttccatgt gaaat
```

//

LOCUS PCU5392 2470 bp DNA PL 24-DEC-1996  
DEFINITION **Pneumocystis carinii** major surface glycoprotein (MSG) gene, partial cds.  
ACCESSION U53921  
NID g1750217  
KEYWORDS  
SOURCE *Pneumocystis carinii*.  
ORGANISM *Pneumocystis carinii*  
Eukaryotae; mitochondrial eukaryotes; Fungi; Ascomycota;  
Archaeascomycetes; Pneumocystidaceae; Pneumocystis.  
REFERENCE 1 (bases 1 to 2470)  
AUTHORS Edman, J.C., Hatton, T.W., Nam, M., Turner, R., Mei, Q., Angus, C.W. and Kovacs, J.A.  
TITLE A single expression site with a conserved leader sequence regulates variation of expression of the *Pneumocystis carinii* family of major surface glycoprotein genes  
JOURNAL DNA Cell Biol. 15 (11), 989-999 (1996)  
MEDLINE 97101113  
REFERENCE 2 (bases 1 to 2470)  
AUTHORS Kovacs, J.A.  
TITLE Direct Submission  
JOURNAL Submitted (09-APR-1996) Joseph A. Kovacs, CCMD/CC, NIH, Bldg. 10, RM. 7D43, MSC1662, Bethesda, MD 20892-1662, USA  
FEATURES  
source Location/Qualifiers  
1..2470  
/organism="Pneumocystis carinii"  
/db\_xref="taxon:4754"  
/lab\_host="rat"  
gene 1940..2470  
/gene="MSG"  
CDS join(<1940..2036,2187..>2470)  
/gene="MSG"  
/note="gpA; expression site of MSG; encodes leader sequence of MSG; surface protein"  
/codon\_start=1  
/product="major surface glycoprotein"  
/db\_xref="PID:g1750218"  
/translation="MRIAFFALFAQLSCILVYSIAERDFMSLDEIYEGGDISFDHEKL  
EFNEYNQVLQMLEKAKKLGTFVDRTKDFSNRRYEGRIELNHLGRRPGVDYFRKGGDV  
FTDGYPRGGHLEDELSEEAAMARP"  
BASE COUNT 707 a 335 c 411 g 1017 t  
ORIGIN  
1 gaattccata ataattgtat tgtaaattat gaatataatt ttgattatct ctttaaaata  
61 tagtattttt ttataataat attttctatt atcaacttga tgttgaaaa atcgtaattt  
121 atcttgatgt ttagagtaaa tgttgatatt ttcaagtcgt attaggtatc cgttggaatt  
181 attttttaat tgaaaaggta taaagtatga ataatatata atatggggaa ttttattttg  
241 tagtatataa atcatagcaa catgattttt ttttgatttt aaatcagatt ttataaaatg  
301 ttttttatat atttataaat cttgttatag tatatatatt caagtgtttt caatgggttt  
361 agtttttaaag agagtgttaa ttaaggtaga aatgttttat taatgggtgat atatatatat  
421 atattgtttt ttgtctaata aaatgaatat tttctaaaaa aaagattatt tttttaaaaa  
481 attgattaaa atgtagaatg tatgattatt tcaagtgatt gtatagtact agtccctcgt  
541 cgtccctcgtc gtcctcgtcg tcctcgtcgt ggtagtgatg agagttttat tgaatcatgt  
601 attttaaaaa aaataaagat tatttatata tatatatatt tgaaaaaatt gaaaggagaa  
661 ttatttagta tgatttttaa attttacaag agttgaaagt tttgttgat ttttatttca  
721 ttattaagggt ttagttggaa cgaagtggaa agtttaatat atgttgaagt agaatatata  
781 gggtaggtgc gaacgttttag gtttaagaag aatatatagg gttacgtcga aggttttaggt  
841 tgaagtagaa tatatagggt taggtctaata gtttaggtta ggtcaaagtg cagggttagg  
901 tctaacattc ggggttaggt caaagtatag gtttaggtcga cgtgtaggtt aggtctgagt  
961 tatgtatatag gtttagggta ggttcactct tttttgaatg aaaaaacttt tttttaataa  
1021 aaaagtaata aaactcaaat ctatagatta acatggatcc gtttttatta taaaatttca  
1081 tgttttatgt ataaaacttt ctttttttta tttttacttg gtccctttct tccctcatc  
1141 atacttgga ttttctgata tattacaata gatggttcat tcctctactc acttttaaaa  
1201 tcttatttaa ctgaccattc atgctatttt tcccctattt tttataaaaa gactcgtcgc  
1261 ttatatatag aatggcgtat acgtaaaaat agaaacgtgt atgcgtgaaa tagaatggtg

```
1321 cttttaaatt tac      tttt attcaaatta ggaaatcttt ttt      tttt tagttttgga
1381 atccattatt cttt      taca aagtatttct ttttttccta aca      tttt ttatacttta
1441 tattttgaaa cctatgcata tactcgatat cactctgact ttgcatattc tactccctcc
1501 tttcctttca cccctctggg gacactcccc tgtccggaca caccocggac agctcttggc
1561 tgtccctccc tgcaccttc tatacccatc tatccatcta tccatctata tatacacctg
1621 ttttactcat ttctatccta taaaaattga cttgtaattt tggatatatt cttctctat
1681 gcatttttag gactggctaa tattttaagg tctagctctt tttatttccg gctagtcac
1741 ccccatccg gaaattggct ttccctttt cagcgttttc ttcgattcat ccgtaaattt
1801 ccgtcaatat ttaaggtct ttgcatatta aggggactgt ttctggactg tatttttacc
1861 ataccctgtg cttccttttag cgcttgaaa atatatttt tcttgatata cgtctcgtct
1921 tcagtttggt tgtgcaataa tgaggattgc attttttgcg ctttttgcg aacttagttg
1981 tatttttagt tattcaatag cagaaaggga ttcatgtca ttagatgaaa tatatggtga
2041 gtttttttcc ctttaatttc tgttgatca ttttttacac gtctttttt actcgtctt
2101 tactcatctt tttactcatc ttctattct ctctcttct tttggaaatc tctttttcct
2161 tttatggttt tactcatgtt ttctagaagg aggcgatata agttttgatc atgaaaaact
2221 cgaatttaac gaatataatc aagttttaca aatgcttgaa aaggcaaaaa aattgggaac
2281 cggctttggt gatagaacca aagatttttc taatagacga tatgaaggga gaattgagtt
2341 aaatcatttg gggagacgcc caggagtcga ctattttagg aaaggtgggg atgtttttac
2401 tgatggttat cctcgtggag gtcatttgat cgaggatgag ttgtccgaag aggcggcaat
2461 ggcacggccg
```

//

LOCUS AF03522. 1945 bp mRNA PLN 02-JAN-1998  
 DEFINITION *Pneumocystis carinii* glycoprotein A (gpA) mRNA, partial cds.  
 ACCESSION AF035226  
 NID g2739185  
 KEYWORDS .  
 SOURCE *Pneumocystis carinii*.  
 ORGANISM *Pneumocystis carinii*  
 Eukaryotae; Fungi; Ascomycota; Archiascomycetes; Pneumocystidaceae;  
*Pneumocystis*.  
 REFERENCE 1 (bases 1 to 1945)  
 AUTHORS Guadiz, G., Haidaris, C.G., Maine, G.N. and Simpson-Haidaris, P.J.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-NOV-1997) Medicine-Vascular Medicine, University of  
 Rochester, 601 Elmwood Avenue, Box 610, Rochester, NY 14642, USA  
 FEATURES Location/Qualifiers  
 source 1..1945  
 /organism="Pneumocystis carinii"  
 /db\_xref="taxon:4754"  
 /tissue\_type="infected lung"  
 /clone="2-1A"  
 gene <1..1945  
 /gene="gpA"  
 CDS <1..1914  
 /gene="gpA"  
 /codon\_start=1  
 /product="glycoprotein A"  
 /db\_xref="PID:g2739186"  
 /translation="KKELRKKNFCNDDSCSNEKDNWEKLVKLRDEDCAKLQSKCFYLE  
 KHCSANLKKPCNNVRVTCYKRGKAAAMTLLSVLKGKLPKGPNNNDYKDCQKALLEKC  
 KEVRNISAIEVFEMCLYPENTCQNLSDIVQQKANYLSLFSLQNRDHPQSQEDCVELQEK  
 EALEADADWLRPPCETLRTHCNFLYLSEKLGKHLLESEGKGLSSNEICNKLNERCHS  
 WHKKKNETYAFPCALRNESCELMVVRVKGKHCDEFKENLKNYNTTLLIKNPSEETCLLWA  
 PHCENLTPNCEEKVAKDINETCRKLQEKCSPVLEKRNKLNKRELKGLTDGKKTEC  
 IDYFKSLCTQADHPNKTALDILCKENGQNIIDDDKAKEKKCQPLIDEVQWECPLLKTKL  
 EKAKDEVAKKADEYKKLKEEAETETKLNVTLPKEAGKATLIRDPKGSSPAKRS  
 VKVEVVSVPITKEHAEIDLLSQALDLYFELREECDALILDCEFKEDCPQCKDSCNSI  
 LATCKGLQPLTIAEPRIEDQKPEAGKEDTPEGGQAGGQAAGKECTSITTSDTWVTST  
 STHTTTTTSTKTTTVKVTLTSTRKCKPTKCTTGKGEPSTRGDEEEVVPSSGGRKWLP  
 LSIVIVGVVALV"

BASE COUNT 744 a 279 c 451 g 471 t

## ORIGIN

```

1 aaaaaagaac ttcgtaaaaa gaatttctgc aatgacgata gctgcagtaa tgaaaaagat
61 aattgggaga aattagtaaa gttaagagat gaagattgtg caaaattaca atctaagtgt
121 ttttatttag aaaaacattg ttcagctaata cttaaaaaaac catgtaacaa tgtagagta
181 acctgttaca agagaggatt aaaggcagcg gctatgactt tgtagagag tggtgtgaaa
241 gggaaactta aaccagggtcc aaataatgat tataaagatt gccaaaaagc attattagaa
301 aagtgtaaag aagtgtgaaa tatcagtgtg gaagtgtgtg agatgtgttt atatccagaa
361 aatacgtgtc aaaatctttc tattgatgtt caacaaaaag caaattattt aagccttttt
421 tcacttcaaa acagggacca tccaagtcaa gaagactgtg ttgaactaca agagaaatgt
481 gaagctttag aagcagatgc agactggctt cgacctcctt gtgaaacatt gagaacgcat
541 tgcaactttc tttacctttc agaaaagttg aaacatcatt tgtagagtga agggaaaagg
601 aaattaagta gcaatgagat ttgcaacaaa gagttgaatg agaggtgtca ctcatggcat
661 aaaaaaaaaga acgaaacgta tgctttttccg tgtgctttgc gcaatgaatc ttgtgagttg
721 atggtttggc gtgtaggaaa acattgtgat gaattttaaag aaaacttgaa aaattataat
781 acaacattaa ttaaaaatcc atcagaagaa acatgtcttt tgtgggcacc acactgtgaa
841 aatctgacgc caaactgtga agaaaagggt gctaaagata tcaatgagac ttgccgtaaa
901 ctccaggaga aatgttcacc agttcttgag aagaggaatt tgaaaaataa attgaaacgt
961 gagttgaaag ggaaattaac ggatggaaag aagacagaat gtatcgatta ttttaaaagt
1021 ctttgtacac aggcagatca tccaaataaa acggcacttg acatattatg taagaaaaat
1081 ggacaaaata ttgatgatga taaagctaaa gaaaagaaat gccaaacctt cattgatgaa
1141 gtacaatggg agtgtccact attaaaaaca aagttagaaa aagcaaagga tgaagtggca
1201 aagaaggcag atgaatataa aaaactaaaa gaagaggcag aaaaggctat tacggaaaca
1261 aaactcaatg ttacagaact accaaaggaa gcaggaaagg caacctgat tagaagagat
1321 ccaaaaggca gcagtccagc caagagatca gtaaaggtag aggttgtgtc tgttccaata

```

```
1381 acgaaagaac atg aggc aattgattta ttatctcaag cat ctct ttatttcgag
1441 ttgagagagg agtg gatgc gttaatatta gattgtgact tta agga ttgtcctcaa
1501 tgtaaagatt catgtaatag tatacttgct acctgtaaag ggttacaacc tcttacaatc
1561 gcggaaccac gtattattga agatcagaaa ccagaagctg gaaaggaaga tactcctgaa
1621 ggaggacaag ctggaggaca agctgcggga aaggagtgc cctctatcac aacaagtgc
1681 acatgggtca cgagtacgtc tacgcatacg acgacgacca cctcgacaaa gacgacgaca
1741 gttaaagtga cgttaacgtc aacgaggaag tgtaagccga ccaagtgtac cactggaaaa
1801 ggagagccga caagtagagg tgatgaagag gaggaggtgg tgccgagtgg aggaaggaag
1861 tggttgcaa gtctaagtat agttattgtt ggtgtggtgg tggcactcgt gtagatcatg
1921 atgaataaga cgttccatgt gaaat
```

//

LOCUS PCU53921 2470 bp DNA PLN. 24-DEC-1996  
 DEFINITION *Pneumocystis carinii* major surface glycoprotein (MSG) gene, partial  
 cds.  
 ACCESSION U53921  
 NID g1750217  
 KEYWORDS .  
 SOURCE *Pneumocystis carinii*.  
 ORGANISM *Pneumocystis carinii*  
 Eukaryotae; mitochondrial eukaryotes; Fungi; Ascomycota;  
 Archaeascomycetes; Pneumocystidaceae; Pneumocystis.  
 REFERENCE 1 (bases 1 to 2470)  
 AUTHORS Edman, J.C., Hatton, T.W., Nam, M., Turner, R., Mei, Q., Angus, C.W. and  
 Kovacs, J.A.  
 TITLE A single expression site with a conserved leader sequence regulates  
 variation of expression of the *Pneumocystis carinii* family of major  
 surface glycoprotein genes  
 JOURNAL DNA Cell Biol. 15 (11), 989-999 (1996)  
 MEDLINE 97101113  
 REFERENCE 2 (bases 1 to 2470)  
 AUTHORS Kovacs, J.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-APR-1996) Joseph A. Kovacs, CCMD/CC, NIH, Bldg. 10,  
 RM. 7D43, MSC1662, Bethesda, MD 20892-1662, USA  
 FEATURES  
 source Location/Qualifiers  
 1..2470  
 /organism="Pneumocystis carinii"  
 /db\_xref="taxon:4754"  
 /lab\_host="rat"  
 gene 1940..2470  
 /gene="MSG"  
 CDS join(<1940..2036,2187..>2470)  
 /gene="MSG"  
 /note="gpA; expression site of MSG; encodes leader  
 sequence of MSG; surface protein"  
 /codon\_start=1  
 /product="major surface glycoprotein"  
 /db\_xref="PID:g1750218"  
 /translation="MRIAFFALFAQLSCILVYSIAERDFMSLDEIYEGGDISFDHEKL  
 EFNEYNQVLQMLEKAKKLGTGFVDRTKDFSNNRYEGRIELNHLGRRPGVDYFRKGGDV  
 FTDGYPRGGHLIEDELSEEAAMARP"  
 BASE COUNT 707 a 335 c 411 g 1017 t  
 ORIGIN  
 1 gaattccata ataattgtat tgtaaattat gaatataatt ttgattatct ctttaaaaata  
 61 tagtattttt ttataataat attttctatt atcaacttga tgttgaaaat atcgtaattt  
 121 atcttgatgt ttagagtaaa tgttgatatt ttcaagtcgt attaggtatc cgttggaatt  
 181 attttttaat tgaaaaggta taaagtatga ataatatata atatggggaa tttatttttg  
 241 tagtatataa atcatagcaa catgattttt ttttgatttt aaatcagatt ttataaaatg  
 301 ttttttatat atttataaat cttgtttatag tatatatatt caagtgtttt caatgggttt  
 361 agtttttaaag agagtgttaa ttaaggtaga aatgttttat taatggtgat atatatatat  
 421 atattgtttt ttgtctaata aaatgaatat tttctaaaat aaagattatt tttttaaaaa  
 481 attgattaaa atgtagaatg tatgattatt tcaagtgtat gtatagtact agtccctcgt  
 541 cgtcctcgtc gtcctcgtcg tcctcgtcgt ggtagtgatg agagttttat tgaatcatgt  
 601 attttaaaaa aaataaagat tatttatata tatatatatt tgaaaaaatt gaaaggagaa  
 661 ttatttagta tgatttttaa attttacaag agttgaaagt tttgttgat tttattttca  
 721 ttattaaggt ttagttggaa cgaagttgga agtttaatat atgttgaagt agaatatata  
 781 gggttaggtc gaacgtttag gtttaagaag aatatatagg gttacgtcga aggttttaggt  
 841 tgaagtagaa tatatagggt taggtctaag gtttaggtta ggtcaaagt cagggtaggt  
 901 tctaaccattc ggggttaggt caaagtatag gtttaggtcga cgtgtaggtt aggtctgagt  
 961 tatgtatagg gttagggtta ggttcactct tttttgaatg aaaaaacttt tttttaataa  
 1021 aaaagtaata aaactcaaat ctatagatta acatggatcc gtttttatta taaaatttca  
 1081 tgttttatgt ataaaacttt ctttttttta tttttacttg gtccctttct ttccctcatc  
 1141 atacttgga tttttctgata tattacaata gatggttcac tcctctactc acttttaaaa  
 1201 tcttatattaa ctgaccattc atgctatttt tcccctattt tttataaaaa gactcgtcgc  
 1261 ttatatatag aatggcgtat acgtaaaaat agaaacgtgt atgcgtgaaa tagaatggtg

```
1321 ctttttaaatt tag      tttt attcaaatta ggaaatcttt ttt      tttt tagttttgga
1381 atccattatt cttt      taca aagtatttct ttttttccta aca      tttt ttatacttta
1441 tatttggaata cctatgcata tactcgatat cactctgact ttgcatattc tactccctcc
1501 tttccttttca cccctcttgg gacactcccc tgcccggaac caccctggac agctcttggc
1561 tgtccctccc  tgcaccttc  tatacccatc tatccatcta tccatctata tatacacctg
1621 ttttactcat  ttctatccta taaaaattga cttgtaattt tggatatatt cttctatat
1681 gcatttttag  gactggctaa ttttttaagg tctagctctt tttatttcct gctagtcac
1741 ccccatccg   gaaattggct tttccttttt cagcgttttc ttcgattcat ccgtaaattt
1801 ccgtcaatat  ttaagggtct ttgcatatta aggggactgt ttctggactg tatttttacc
1861 ataccctgtg  cttccttttag cgcttggaata atatattttt tcttgatata cgtctcgtct
1921 tcagtttggt  tgtgcaataa tgaggattgc attttttgcg ctttttgcg  aacttagttg
1981 tatttttagtt tattcaatag cagaaaggga tttcatgtca ttagatgaaa tatatggtga
2041 gttttttttc  ctttaatttc  tgttgatatc ttttttacac gtcttttttt actcgtcttt
2101 tactcatctt  tttactcatc tttctattct ctctcttct  tttggaaatc tcttttcct
2161 tttatgggtt  tactcatggt ttctagaagg aggcgatata agttttgatc atgaaaaact
2221 cgaatttaac  gaatataatc aagttttaca aatgcttgaa aaggcaaaaa aattgggaac
2281 cggctttgtt  gatagaacca aagatttttc taatagacga tatgaaggga gaattgagtt
2341 aaatcatttg  gggagacgcc caggagtcca ctattttagg aaaggtgggg atgtttttac
2401 tgatgggttat ctcgtggag  gtcatttgat cgaggatgag ttgtccgaag aggcggcaat
2461 ggcacggccg
```

//

LOCUS PCU3966 629 bp mRNA 29-NOV-1995  
DEFINITION *Pneumocystis carinii* major surface glycoprotein (msg) mRNA, partial cds.  
ACCESSION U39660  
NID g1079706  
KEYWORDS .  
SOURCE *Pneumocystis carinii*.  
ORGANISM *Pneumocystis carinii*  
Eukaryotae; mitochondrial eukaryotes; Fungi; Ascomycota;  
Archaeascomycetes; Pneumocystidaceae; *Pneumocystis*.  
REFERENCE 1 (bases 1 to 629)  
AUTHORS Kovacs, J.A., Edman, J.C. and Angus, C.W.  
TITLE Direct Submission  
JOURNAL Submitted (30-OCT-1995) J.A. Kovacs, CCMD, NIH, Building 10, Rm 7D43, MSC 1662, Bethesda, MD 20892-1662, USA

FEATURES  
source Location/Qualifiers  
1..629  
/organism="Pneumocystis carinii"  
/db\_xref="taxon:4754"  
gene 55..629  
/gene="msg"  
CDS 55..>629  
/gene="msg"  
/codon\_start=1  
/product="major surface glycoprotein"  
/db\_xref="PID:g1079707"  
/translation="MRIAFFALFAQLSCILVYSIAERDFMSLDEIYGGDISFDHEKLE  
FNEYNQVLQMLEKAKKLGTGFVDRTKDFSNRRYEGRIELNHLGRRPGVDYFRKGGDVF  
TDGYPRGGHLIEDELSEEAAMARPVKQAVQGAQDEIDEKHLAFIVKDKYKEEQKCK  
EELEKYCKELKEADKNLENVDDKVKGLCDDK"  
misc\_difference 228..234  
/gene="msg"  
/note="in some cDNA clones, there are 6 A's at this  
position, instead of 7. This appears to be due to an error  
in reverse transcription"  
/replace="aaaaaa"

BASE COUNT 218 a 73 c 154 g 184 t  
ORIGIN

```
1 ggaaaatata ttttttcttg atatccgtct cgtcttcagt ttgtttgtgc aataatgagg
61 attgcatttt ttgcgctttt tgcgcaactt agttgtattt tagtttattc aatagcagaa
121 agggatttca tgtcattaga tgaaatatat ggaggcgata taagttttga tcatgaaaaa
181 ctcgaaattta acgaatataa tcaagtttta caaatgcttg aaaaggcaaa aaaattggga
241 accggccttg ttgatagaac caaagatttt tctaatagac gatatgaagg gagaattgag
301 ttaaatcatt tggggagacg cccaggagtc gactatttta ggaaaggtag ggatgttttt
361 actgatgggt atcctcgtgg aggtcatttg atcgaggatg agttgtccga agaggcggca
421 atggcacggc cggttaagag gcaagcagta caaggagcac aagatgagat tgatgagaaa
481 caccttttgg ctttcattgt gaaggacaaa tataaagaag acaaaaaatg caaagaagaa
541 ctcgagaaat attgtaaaga gttgaaggaa gcagataaaa atctagagaa tgtggatgat
601 aaagttaaag gactttgtga tgataaaaa
```

//



LOCUS AF043102 4002 bp mRNA PLN 06-JUN-1998  
DEFINITION *Pneumocystis carinii* surface glycoprotein A (gpA) mRNA, complete cds.  
ACCESSION AF043102  
NID g3184385  
KEYWORDS .  
SOURCE *Pneumocystis carinii*.  
ORGANISM *Pneumocystis carinii*  
Eukaryota; Fungi; Ascomycota; Archiascomycetes; Pneumocystidaceae; *Pneumocystis*.  
REFERENCE 1 (bases 1 to 4002)  
AUTHORS Haidaris, C.G., Medzihradsky, O.F., Gigliotti, F. and Simpson-Haidaris, P.J.  
TITLE Molecular characterization of mouse *Pneumocystis carinii* surface glycoprotein A  
JOURNAL DNA Res. 5, 77-85 (1998)  
REFERENCE 2 (bases 1 to 4002)  
AUTHORS Haidaris, C.G., Medzihradsky, O.F., Gigliotti, F. and Simpson-Haidaris, P.J.  
TITLE Direct Submission  
JOURNAL Submitted (14-JAN-1998) Microbiol/Immunol, U. Rochester Med. Ctr., Box 672, 601 Elmwood Ave., Rochester, NY 14642, USA  
FEATURES  
source Location/Qualifiers  
1..4002  
/organism="Pneumocystis carinii"  
/db\_xref="taxon:4754"  
/lab\_host="C.B-17 SCID mouse"  
gene 1..4002  
/gene="gpA"  
CDS 52..3849  
/gene="gpA"  
/codon\_start=1  
/product="surface glycoprotein A"  
/db\_xref="PID:g3184386"  
/translation="MRIAFFALFAQLSYVLGSFLRDRDLPNEEDVYGYENFGLDPNSP  
EPSEFLNIIEMLRNAKRLGSGKINQGQFLSNRKARRDFDLCRSCNRPGVDYFRKSDYD  
GFFSEDFSSSEDYSQDKRWVEEVAQKEAAMAQPVKRQAAGQAAGNDEIKEEQVLGLIVK  
SGYNNDNKCKANLKHCEELKKIDGKLESVDVKVGLCENGKEGEKCKELKKKLETEL  
GAFKTEVENALNNTDEKCRKYEEKCLLLEEGDPNNLEEKCVKLRDRCYRQRQGVAK  
EILLRALEGKVNNDCKKRMKEICQGLSEYSDELVFSCFNSDKTCEYLQKNHGDSCK  
PLEKELEDKELVEKCEYLEKCYFYGSSCKDTKCDKVNKCKGKGIEYEGPKLDFSPV  
REKPRFPKEIEVENLYKKEEAKGIIVGPKPKYKTLRDLALLIKERNKGDEGEKCKKAL  
EDCESFKHLDYGLEELCGDKDKEDRKCELVEVEDRCTNFKLELYLKGLSTEFKDKES  
DYFSWGQVSKLVSMEDCIKFSECFHLERVCTNKIGKACENVRVACYKKGQDRVLNRY  
FQEGKGLIGDLELVTENLEKCKQSVVGNITKLKEDRRYFTKCHLPKLCYELDDVI  
LQSEELVVLNLRDFPRKEDCVELKKKCKDLES DSYLNHEKCDTLNRRCEYLKVTEE  
LRKRLKRGDDALRTQGNCTAVLKKECEELSKRGKDEFSVSCALREETCSFMVEQTEN  
ECLFLKNMENGKIINKIEKGNELTVEELCTLFDPYCHQYIENCPRDLKKASNSNKNK  
VCLELEEKCKPFFEKLKLENELTHELKGSLDKDDECKKALRKHCSEQKNSANQKFNSL  
CNTDKDKDVEEKVKCKLVEKVKRKCPTELENKLNEEKNELKKKKDEYEKAKQSEKFTK  
EAKLVLSRPEQDGGGSKAQDGSVPKVPVGPVPPAPAPQPTGGVPAPTLPAPAPQPT  
SGGAPLPVPPAPAPAPGAPSTPGTPAAPAGPAPAGTPSTPSTPPAGPAGPSGGTPGAP  
AGPPAPGGSTPSGTTNTSNVILVRRFTVSGEVSEPEKKAFVATARALELYLELKEKCK  
GLKGDCEFRKDCPKCETVCKEIDELCEGIEGLKVTPHHTVTSTATQT TTTTATTTTT  
TTTTTTTTTTTTTTATTTESVDGGKVTEECTLVOTTDTWVTSTSLHTSTLTSTSTVTST  
VTLTSMRKCKPTRCTSDSSKETETQKEEKEEVEKPNEGMKIRVPEMIKIMLLGVIVM  
GML"  
BASE COUNT 1616 a 487 c 940 g 959 t  
ORIGIN  
1 caaaaatata tttttcttga tatcogtctc ttttctcaa tttgtgcaat tatgaggatt  
61 gcattttttg cgctttttgc gcaacttagt tatgtcttgg ggagtttttt aagagataga  
121 gatttgccaa atgaggaaga tgtttatggt tatgaaaatt ttgggttgga tcccaattca  
181 ccggaaccga gtgaattttt aaatattata gaaatgcttc gaaatgcaa gcgattagga  
241 agtgggaaga ttaaccaggg tcaatttctc tcgaatcgta aggctaggag agattttgat

LOCUS D31920 81 bp mRNA PM 22-SEP-1997  
DEFINITION *Pneumocystis carinii* MSG mRNA for major surface glycoprotein,  
5'UTR.  
ACCESSION D31920  
NID g2443384  
KEYWORDS MSG.  
SOURCE *Pneumocystis carinii* cDNA to mRNA.  
ORGANISM *Pneumocystis carinii*  
Eukaryotae; Fungi; Ascomycota; Archaeascomycetes;  
Pneumocystidaceae; *Pneumocystis*.  
REFERENCE 1 (sites)  
AUTHORS Wada,M., Sunkin,S.M., Stringer,J.R. and Nakamura,Y.  
TITLE Antigenic variation by positional control of major surface  
glycoprotein gene expression in *Pneumocystis carinii*  
J. Infect. Dis. 171 (6), 1563-1568 (1995)  
JOURNAL  
MEDLINE 95287050  
REFERENCE 2 (bases 1 to 81)  
AUTHORS Nakamura,Y.  
TITLE Direct Submission  
JOURNAL Submitted (24-JUN-1994) to the DDBJ/EMBL/GenBank databases.  
Yoshikazu Nakamura, Institute of Medical Science, University of  
Tokyo, Tumor Biology; Shirokanedai 4-6-1, Minato-ku, Tokyo 108,  
Japan (E-mail:nak@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5307,  
Fax:03-5449-5415)  
FEATURES Location/Qualifiers  
source 1..81  
/organism="*Pneumocystis carinii*"  
/db\_xref="taxon:4754"  
gene 1..81  
/gene="MSG"  
5'UTR 1..81  
/gene="MSG"  
/note="major surface glycoprotein"  
BASE COUNT 30 a 10 c 13 g 28 t  
ORIGIN  
1 attcatgaag taactatcaa gcttaataaa ggaccaata taggagtata ttgttaatat  
61 ttatgctata tattgggatac c  
//

LOCUS D31914 496 bp mRNA 22-SEP-1997  
 DEFINITION *Pneumocystis carinii* MSG mRNA for major surface glycoprotein,  
 5'UTR.  
 ACCESSION D31914  
 NID g2443378  
 KEYWORDS MSG.  
 SOURCE *Pneumocystis carinii* cDNA to mRNA.  
 ORGANISM *Pneumocystis carinii*  
 Eukaryotae; Fungi; Ascomycota; Archaeascomycetes;  
 Pneumocystidaceae; *Pneumocystis*.  
 REFERENCE 1 (sites)  
 AUTHORS Wada,M., Sunkin,S.M., Stringer,J.R. and Nakamura,Y.  
 TITLE Antigenic variation by positional control of major surface  
 glycoprotein gene expression in *Pneumocystis carinii*  
 J. Infect. Dis. 171 (6), 1563-1568 (1995)  
 MEDLINE 95287050  
 REFERENCE 2 (bases 1 to 496)  
 AUTHORS Nakamura,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-JUN-1994) to the DDBJ/EMBL/GenBank databases.  
 Yoshikazu Nakamura, Institute of Medical Science, University of  
 Tokyo, Tumor Biology; Shirokanedai 4-6-1, Minato-ku, Tokyo 108,  
 Japan (E-mail:nak@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5307,  
 Fax:03-5449-5415)  
 FEATURES Location/Qualifiers  
 source 1..496  
 /organism="Pneumocystis carinii"  
 /db\_xref="taxon:4754"  
 gene 1..496  
 /gene="MSG"  
 5'UTR 1..496  
 /gene="MSG"  
 /note="major surface glycoprotein"  
 BASE COUNT 214 a 50 c 89 g 143 t  
 ORIGIN  
 1 caactatcaa gcttaataaa ggaccaata taggagtata ttgttaatat ttatgctata  
 61 tattgggatc cagaaataat atgaaatgct tttttccaaa attaataaaa ttatagttct  
 121 agaataatta attttaattg ttcctactta tggcaatcaa agcacacatt tgaaagccag  
 181 agattctcaa gtaagcgag caaatactcc gttaaggat attgttgta aggatgaaga  
 241 aatccttgct tatattttga gagaagatta taaagataaa tgcgaactaa agcttgagga  
 301 atattgtaaa gaattgaggg atatagatca aggattaaat aaagttcata ctatagttaa  
 361 agaaatttgt gatgatgaaa aacgagacgg aaaatgcaaa gaactgaaag acaaagttaa  
 421 aaaggaattg gaaactttta aagaggaact tgaaaaagca ttgaaagaca taaaagatga  
 481 aaattgtgaa aaatat

//